Wherefore, what is claimed is:

1. A computer-implemented process for self calibrating a plurality of audio sensors of a microphone array, wherein each sensor has a known location and generates a signal representing a channel of the array, said process comprising using a computer to perform the following process actions:

inputting a set of substantially contemporaneous audio frames extracted from the signals generated by at least two sensors of the array and a direction of arrival (DOA) associated with the frame set;

computing the energy of each frame;

establishing an approximation function that characterizes the relationship between the locations of the sensors and their computed energy values and using the function to estimate the energy of each frame; and

for each frame, computing an estimated gain that compensates for the difference between the computed energy of the frame and its estimated energy, and applying the gain to the next frame associated with the same audio sensor.

2. The process of Claim 1, wherein the process action of inputting the set of audio frames, comprises an action of inputting the audio frames and associated DOA only if the frames comprise audio data exhibiting evidence of a single dominant sound source.

3. The process of Claim 1, wherein the process action of establishing the approximation function, comprises the actions of:

projecting the location of each sensor associated with an input frame onto a line defined by the DOA;

establishing the straight line function that characterizes the relationship between the projected locations of the sensors on the DOA line and the computed energy values of the frames associated with the sensors; and

estimating the energy of each frame using the straight line function.

4. The process of Claim 3, wherein the process action of projecting the location of each sensor associated with an input frame onto a line defined by the DOA, comprises an action of projecting the locations of the sensors, which are known in terms of a radial coordinate system with the centroid of the microphone array as its origin, onto the DOA line.

5. The process of Claim 1, further comprising a process action of normalizing the computed gain estimates by dividing each by the average of all the gain estimates.

6. The process of Claim 1, further comprising inputting a series of substantially contemporaneous audio frame sets extracted from the signals generated by at least two sensors of the array and a DOA associated with each frame set, wherein the audio frames are input only if they comprise audio data exhibiting evidence of a single dominant sound source, and repeating the process actions of Claim 1 for each set of frames input.

7. The process of Claim 6, wherein the number of sets of substantially contemporaneous audio frames input over a prescribed time period is limited to a prescribed number to reduce computational costs.

8. The process of Claim 6, further comprising a process action of adaptively refining the gain each time a gain is computed, said refining action comprising:

establishing an adaptation parameter that dictates the weight a currently computed gain is given; and

computing the refined gain as the sum of the gain multiplied by the adaptation parameter, and a refined gain computed for the immediately preceding frame input from of the same array channel as the frame used to

compute the gain under consideration multiplied by one minus the adaptation parameter.

9. The process of Claim 8, wherein the adaptation parameter is selected within a range of parameter values between about 0.001 and about 0.01.

10. The process of Claim 9, wherein an adaptation parameter closer to 0.01 is chosen if calibrating a microphone array operated in a controlled environment wherein reverberations are minimal.

11. The process of Claim 9, wherein an adaptation parameter closer to 0.001 is chosen if calibrating a microphone array operated in an environment wherein reverberations are not minimal.

12. The process of Claim 8, further comprising the process actions of:

monitoring the value of each refined gain computed for a channel of the array;

determining if the difference between the values of a prescribed number of consecutively computed refined gains exceeds a prescribed change threshold;

whenever it is found that the change threshold is not exceeded, suspending the inputting of any further frames associated with the affected channel of the array.

13. The process of Claim 12, further comprising, whenever the inputting of further frames has been suspended for an array channel, performing the process actions of:

periodically inputting at least one new audio frame extracted from the signal generated by the sensor of the array associated with the array channel

under consideration, wherein the audio frame is input only if it comprises audio data exhibiting evidence of a single dominant sound source;

determining if the difference between the last, previously-computed refined gain for the channel and the current gain computed for the channel exceeds the prescribed change threshold; and

whenever it is found that the change threshold is exceeded, reinitiating the inputting of further frame sets.

14. A system for self calibrating the audio sensors of a microphone array, comprising:

a microphone array having a plurality of audio sensors generating signals each of which represents a channel of the array;

a general purpose computing device;

a computer program comprising program modules executable by the computing device, wherein the computing device is directed by the program modules of the computer program to,

input a set of substantially contemporaneous audio frames extracted from the signals generated by at least two sensors of the array, wherein the audio frames are input only if they comprise audio data exhibiting evidence of a single dominant sound source,

input a direction of arrival (DOA) associated with inputted the frames,

for each set of frames and associated DOA input,

compute the energy of each frame,

project a pre-established location of each sensor associated with an input frame onto a line defined by the DOA

establish an approximation function that characterizes the relationship between the projected locations of the sensors on the DOA line and the computed energy values of the frames associated with the sensors,

estimate the energy of each frame using the approximation function,

for each frame, compute an estimated gain that compensates for the difference between the computed energy of the frame and its estimated energy,

normalize the computed gain estimates by dividing each by the average of the gain estimates, and

respectively apply each of the normalized gain estimates to the next frame associated with the same sensor.

15. The system of Claim 14, wherein the program module for computing the energy of each frame, comprises a sub-module for computing $E_m = \frac{1}{N}\sum_{k=0}^{N-1} b_m(kT)^2$, where E_m is the computed energy of the frame of the m^{th} sensor, N is the number of samples associated with the inputted audio frame under consideration, $b_m(kT)$ is the input sample from the m-th sensor at moment kT, and T is the sampling period used to generate the frames.

16. The system of Claim 14, wherein the program module for projecting the pre-established location of each sensor associated with an input frame onto the line defined by the DOA, comprises a sub-module for projecting the locations of the sensors, which are known in terms of a radial coordinate system with the centroid of the microphone array as its origin, onto the DOA line.

17. The system of Claim 14, wherein the program module for establishing an approximation function that characterizes the relationship between the projected locations of the sensors on the DOA line and the computed energy values associated with the sensors, comprises sub-modules for:

defining a straight line function as having the form $\tilde{E}(d) = a_1 d + a_0$, wherein $\tilde{E}(d)$ is the estimated energy of a frame, d is the projected location of the sensor associated with the frame, and a_1 and a_0 unknown coefficients;

computing the values of a_1 and a_0 that produce estimated energy values for each projected sensor location that satisfy the Least Means Squares requirement such that $\left(\sum_{i=0}^{M-1}(\tilde{E}(d_i)-E_i)^2\right)$ is minimized where M is the number of sensors having an inputted frame associated therewith and E is the computed energy of a frame.

18. The system of Claim 17, wherein the program module for establishing an approximation function further comprises sub-modules for, whenever the coefficient a_1 is computed to be less than zero:

setting the coefficient a_1 to zero; and

setting the coefficient a_0 to the average of the computed energy values associated with the sensors.

19. The system of Claim 17, wherein the program module for computing an estimated gain that compensates for the difference between the computed energy of the frame and its estimated energy, comprises a sub-module for computing $g_m = G_m^{n-1}\sqrt{\frac{E_m}{\tilde{E}(d_m)}}$,where g_m is the estimated gain, and where G_m^{n-1} is the last gain computed for the channel under consideration or 1 if the gain has not been computed before.

20. The system of Claim 14, further comprising a program module for discarding the normalized gains computed the set of frames under consideration whenever the estimated gain of the current frame is outside a prescribed range of acceptable gain values.

21. The system of Claim 20, wherein the prescribed range of acceptable gain values comprises gain values ranging from about 0.5 to about 2.0.

22. The system of Claim 19, wherein the program module for respectively applying each of the normalized gain estimates to the frame associated with the same sensor, comprises a sub-module for multiplying the frame by the gain estimate associated with the array channel where the frame was extracted.

23. The system of Claim 14, further comprising a program module for adaptively refining the normalized gain for each sensor, said refining module comprising sub-modules for:

establishing an adaptation parameter that dictates the weight a currently computed normalized gain is given;

computing the refined normalized gain as $G_m^n = (1-\alpha)G_m^{n-1} + \alpha G_m$, where G_m^n is the refined normalized gain, G_m^{n-1} is the last previously-computed refined normalized gain for the same array channel, and α is the adaptation parameter.

24. The system of Claim 23, wherein the adaptation parameter is selected within a range of parameter values between about 0.001 and about 0.01, and wherein an adaptation parameter closer to 0.01 is chosen if calibrating a microphone array operated in a controlled environment wherein reverberations are minimal, and wherein an adaptation parameter closer to 0.001 is chosen if calibrating a microphone array operated in an environment wherein reverberations are not minimal.

25. The system of Claim 23, further comprising program modules for:

monitoring the value of each refined normalized gain computed for a channel of the array;

determining if the difference between the values of consecutively computed refined normalized gains in any channel exceeds a prescribed change threshold within a prescribed period of time;

whenever it is found that the change threshold is not exceeded in any channel, suspending the inputting of any further frame sets.

26. The system of Claim 25, further comprising program modules for, whenever the inputting of further frames sets has been suspended:

periodically inputting at least one new audio frame set, wherein the audio frame set is input only if the frames comprise audio data exhibiting evidence of a single dominant sound source;

computing normalized gain estimates for the set;

determining if the difference between the last, previously-computed refined normalized gain for any channel and the current normalized gain computed for channel the exceeds the prescribed change threshold; and

whenever it is found that the change threshold is exceeded, reinitiating the inputting of further frame sets.

27. A computer-readable medium having computer-executable instructions for self calibrating a plurality of audio sensors of a microphone array, wherein each sensor has a known location and generates a signal representing a channel of the array, said computer-executable instructions comprising:

inputting a series of substantially contemporaneous audio frame sets extracted from the signals generated by at least two sensors of the array and a direction of arrival (DOA) associated with each frame set, wherein an audio frame set is input only if the frames thereof comprise audio data exhibiting evidence of a single dominant sound source;

for each frame set inputted,

computing the energy of each frame,

establishing an approximation function that characterizes the relationship between the locations of the sensors and their computed energy values and using the function to estimate the energy of each frame, and

for each frame, computing an estimated gain that compensates for the difference between the computed energy of the frame and its estimated energy, and applying the gain to the frame.

28. The computer-readable medium of Claim 27, wherein the instruction for establishing the approximation function, comprises sub-instructions for:

projecting the location of each sensor associated with an input frame onto a line defined by the DOA;

establishing a straight line function that characterizes the relationship between the projected locations of the sensors on the DOA line and the computed energy values of the frames associated with the sensors; and

estimating the energy of each frame using the straight line function.

29. The computer-readable medium of Claim 28, further comprising an instruction for normalizing the computed gain estimates by dividing each by the average of all the gain estimates.

30. The computer-readable medium of Claim 29, further comprising an instruction for adaptively refining the normalized gain each time a gain is computed, said refining instruction comprising sub-instructions for:

establishing an adaptation parameter that dictates the weight a currently computed normalized gain is given; and

computing the refined normalized gain as the sum of the normalized gain multiplied by the adaptation parameter, and a refined normalized gain computed for the immediately preceding frame input from of the same array channel as the frame used to compute the normalized gain under consideration multiplied by one minus the adaptation parameter.

31. The computer-readable medium of Claim 30, wherein the sub-instruction for establishing an adaptation parameter, comprises selecting the

adaptation parameter to be within a range of parameter values between about 0.001 and about 0.01, and wherein an adaptation parameter closer to 0.01 is chosen if calibrating a microphone array operated in a controlled environment wherein reverberations are minimal, and wherein an adaptation parameter closer to 0.001 is chosen if calibrating a microphone array operated in an environment wherein reverberations are not minimal.